CLAIMS

What is claimed is:

1. A method of automatically interconnecting a new node into a cluster of other nodes, each node including a plurality of processor units and an internal switching fabric, one of the processor units in each node hosting a monitor process, comprising:

(a) broadcasting a discovery probe packet from the monitor process in the new node to each of the plurality of processor units of all the other nodes in the cluster, the discovery probe packet containing configuration information about the new node;

(b) receiving at least one discovery probe packet at the monitor process of a discovery initiator node, a discovery initiator node being any one of the other nodes of the cluster;

(c) setting up, at the discovery initiator node, in response to the discovery probe packet, connection information enabling the new node to directly communicate with the monitor process of the discovery initiator node;

(d) sending a discovery initiator packet to all of the plurality of processor units in the new node, the discovery initiator packet containing the connection information set up at the discovery initiator node for the new node;

(e) receiving at least one discovery initiator packet at the monitor process of the new node;

(f) setting up, at the new node, in response to the discovery initiator packet, connection information enabling the discovery initiator node to directly communicate with the monitor process of the new node;

(g) sending a discovery confirmation packet directly to the monitor process of the discovery initiator node, the discovery confirmation packet containing the connection information set up at the new node for the discovery initiator node;

(h) receiving the discovery confirmation packet at the monitor process of the discovery initiator node;

(j) sending, in response to the received discovery confirmation packet, a discovery acknowledgement packet directly to the monitor process of the new node; and

(k) receiving at the monitor process of the new node the discovery acknowledgment packet.

2. A method of automatically interconnecting a new node as recited in claim 1, wherein the connection information includes interrupt and barrier addresses.

3. A method of automatically interconnecting a new node as recited in claim 1,

wherein each processor unit of a node has a message receiving process; and

wherein the step of receiving at least one discovery probe packet at the monitor process of a discovery initiator node includes:

(l) receiving at the message receiving processes of the discovery initiator node the discovery probe packets; and

(m) forwarding the discovery probe packets from the message receiving processes to the monitor process of the discovery initiator node.

4. A method of automatically interconnecting a new node as recited in claim 3, wherein the step of forwarding is performed by means of an interprocess communications message.

5. A method of automatically interconnecting a new node as recited in claim 3, wherein the step of receiving the discovery probe packets includes responding to a permissive interrupt in the processor unit of the message receiving process.

6. A method of automatically interconnecting a new node as recited in claim 5, wherein responding to a permissive interrupt in the processor unit of the message receiving process includes:

(n) marking a well-known entry in an access validation and translation table with access permissions for all processor units in the cluster;

(p) receiving packets at the well-known entry without performing access permission checks; and

(q) raising a permissive interrupt in response to the received packet.

7. A method of automatically interconnecting a new node as recited in claim 1,

wherein each processor unit of a node has a message receiving process; and

wherein the step of receiving at least one discovery initiator packet at the monitor process of the new node includes:

(l) receiving at the message receiving processes of the new node the discovery initiator packets; and

(m) forwarding the discovery initiator packets from the message receiving process to the monitor process of the new node.

8. A method of automatically interconnecting a new node as recited in claim 7, wherein the step of forwarding is performed by means of an interprocess communications message.

9. A method of automatically interconnecting a new node as recited in claim 7, wherein the step of receiving the discovery initiator packets includes responding to a permissive interrupt in the processor unit of the message receiving process.

10. A method of automatically interconnecting a new node as recited in claim 1, further comprising the step of installing logical network devices in the discovery initiator node for all of the processor units in the new node after receiving the discovery probe packet, if the discovery probe packet is valid.

11. A method of automatically interconnecting a new node as recited in claim 1, further comprising the step of installing a logical network device in the new node for the processing unit hosting the monitoring process after receiving the discovery initiator packet, if the discovery initiator packet is valid.

12. A method of automatically interconnecting a new node as recited in claim 1, further comprising:

(l) verifying the received discovery probe packet; and

(m) if the discovery probe packet has an error, logging that an error occurred at the broadcast stage.

13. A method of automatically interconnecting a new node as recited in claim 1, further comprising:

(l) verifying the received discovery initiator packet; and

(m) if the discovery initiator packet has an error, logging that an error occurred at the initiation stage.

14. A method of automatically interconnecting a new node as recited in claim 1, further comprising:

(l) verifying the received discovery confirmation packet; and

(m) if the discovery confirmation packet has an error, logging that an error occurred at the confirmation stage.

15. A method of automatically interconnecting a new node as recited in claim 1, further comprising:

(l) verifying the received discovery acknowledgment packet; and

(m) if the discovery acknowledgment packet has an error, logging that an error occurred at the acknowledgment stage.

16. A method of automatically interconnecting a new node into a cluster of other nodes, each node including a plurality of processor units and an internal switching fabric, one of the processor units in each node hosting a monitor process, comprising:

(a) broadcasting a discovery probe packet from the monitor process in the new node to each of the plurality of processor units of all the other nodes in the cluster, the discovery probe packet containing configuration information about the new node;

(b) receiving at least one discovery probe packet at the monitor process of a discovery initiator node, the discovery initiator node being one of the other nodes of the cluster;

(c) verifying the received discovery probe packet;

(i) if the discovery probe packet is valid,

(d) setting up, at the discovery initiator node, connection information enabling the new node to directly communicate with the monitor process of the discovery initiator node; and

(e) sending a discovery initiator packet to all of the plurality of processor units in the new node, the discovery initiator packet containing the connection information set up at the discovery initiator node for the new node;

(f) receiving at least one discovery initiator packet at the monitor process of the new node;

(g) verifying the received discovery initiator packet;

(ii) if the discovery initiator packet is valid,

(h) setting up, at the new node, connection information enabling the discovery initiator node to directly communicate with the monitor process of the new node;

(j) sending a discovery confirmation packet directly to the monitor process of the discovery initiator node, the discovery confirmation packet containing the connection information set up at the new node for the discovery initiator node;

(l) verifying the discovery confirmation packet at the monitor process of the discovery initiator node; and

(iii) if the discovery confirmation packet is valid,

(l) sending a discovery acknowledgement packet directly to the monitor process of the new node.

17. A method of automatically interconnecting a new node into a cluster of other nodes, each node including a plurality of processor units and an internal switching fabric, one of the processor units in each node hosting a monitor process, comprising:

(a) broadcasting a discovery probe packet from the monitor process in the new node to each of the plurality of processor units of all the other nodes in the cluster, the discovery probe packet containing configuration information about the new node;

(b) receiving at least one discovery initiator packet at the monitor process of the new node, the discovery initiator packet containing connection information enabling the new node to directly communicate with the monitor process of at least one other node as the discovery initiator node;

(c) setting up, at the new node, in response to the discovery initiator packet, connection information enabling the discovery initiator node to directly communicate with the monitor process of the new node;

(d) sending a discovery confirmation packet directly to the monitor process of the discovery initiator node, the discovery confirmation packet containing the connection information set up at the new node for the discovery initiator node; and

(e) receiving a discovery acknowledgement packet in response to the discovery confirmation packet.

18. A method of automatically interconnecting a new node into a cluster of other nodes as recited in claim 17,

wherein processor units in the cluster have processor unit ids;

further comprising, prior to broadcasting a discovery probe packet,

(f) installing a discovery broadcast device (DBD) which is used for broadcasting packets to all other processor units of nodes in the cluster; and

(g) reconfiguring the discovery broadcast device to temporarily assign a destination processor unit id to the discovery broadcast device; and

wherein the step of broadcasting a discovery probe packet includes broadcasting a discovery probe packet by the discovery broadcast device.

19. A method of automatically interconnecting a new node into a cluster of other nodes as recited in claim 17,

wherein there is queue for discovery probe broadcast packets to be sent by the new node; and

wherein a number of packets in the queue is limited.

20. A method of automatically interconnecting a new node into a cluster of other nodes as recited in claim 17,

further comprising, prior to broadcasting a discovery probe packet,

queuing a discovery probe packet;

requesting a transfer completion interrupt;

suspending program execution of the monitor process;

further comprising, after broadcasting the discovery probe packet,

receiving a transfer completion interrupt; and

resuming execution of the monitor process.

21. A method of automatically interconnecting a new node into a cluster of other nodes as recited in claim 17,

further comprising, prior to broadcasting a discovery probe packet,

queuing a discovery probe packet;

requesting a transfer completion interrupt;

suspending program execution of the monitor process;

further comprising, after broadcasting the discovery probe packet,

receiving a timeout error interrupt; and

resuming execution of the monitor process.

22. A method of automatically interconnecting a new node into a cluster of other nodes as recited in claim 21, wherein a time period between broadcasting the discovery probe packet and the timeout error interrupt permits other packets to be interleaved with broadcast discovery probe packets.

23. A method of automatically interconnecting a new node into a cluster of other nodes, each node including a plurality of processor units, each having a processor number, and an internal switching fabric, one of the processor units in each node hosting a monitor process, comprising:

(a) broadcasting a discovery probe packet from the monitor process in the new node to processor units of nodes in the cluster, wherein the processor numbers of the processor units are equal and the discovery probe packet contain configuration information about the new node;

(b) receiving at least one discovery initiator packet at the monitor process of the new node, the discovery initiator packet containing connection information enabling the new node to directly communicate with the monitor process of at least one other node as the discovery initiator node;

(c) setting up, at the new node, in response to the discovery initiator packet, connection information enabling the discovery initiator node to directly communicate with the monitor process of the new node;

(d) sending a discovery confirmation packet directly to the monitor process of the discovery initiator node, the discovery confirmation packet containing the connection information set up at the new node for the discovery initiator node; and

(e) receiving a discovery acknowledgement packet in response to the discovery confirmation packet.

24. A method of automatically interconnecting a new node into a cluster of other nodes, each node including a plurality of processor units as recited in claim 23,

wherein there is a minimum processor number and a maximum processor number and a current processor number is set at the minimum processor number;

wherein the broadcasting step includes broadcasting with a processor number equal to the current processor number; and

further comprising;

(f) incrementing the current processor number after broadcasting the discovery probe packets to the processor units with the current processor number in all other nodes that have not yet replied with a discovery initiation packet; and

(g) if the current processor number is not more than a maximum processor number, continuing at the broadcasting step with the current processor number.

25. A method of automatically interconnecting a new node into a cluster of other nodes, each node including a plurality of processor units as recited in claim 24, wherein a time period of about one second or less occurs between the broadcasting steps with consecutive processor numbers.

26. A method of automatically interconnecting a new node into a cluster of other nodes, each node including a plurality of processor units as recited in claim 24,

wherein a subset of processor numbers for other nodes is known to be valid; and

wherein, if the current processor number is more than the maximum processor number, continuing at the broadcasting step using only the subset of valid processor numbers for other nodes from which no discovery initiator packet was received.

27. A method of automatically interconnecting a new node into a cluster of other nodes, each node including a plurality of processor units as recited in claim 26, wherein a time period of about 10 seconds occurs between broadcasting steps using only the subset of valid processor numbers for other nodes from which no discovery initiator packet was received.

28. A method of automatically interconnecting a new node into a cluster of other nodes, each node including a plurality of processor units as recited in claim 26, wherein only processor numbers that are likely to be present in other nodes from which no discovery initiator packet was received are used in the broadcasting step.

29. A method of automatically interconnecting a new node into a cluster of other nodes, each node including a plurality of processor units and an internal switching fabric, one of the processor units in each node hosting a monitor process, comprising:

(a) receiving at least one discovery probe packet at the monitor process of a discovery initiator node, the discovery initiator node being one of the other nodes of the cluster, the discovery probe packet containing configuration information about the new node;

(b) setting up, at the discovery initiator node, in response to the discovery probe packet, connection information enabling the new node to directly communicate with the monitor process of the discovery initiator node;

(c) sending a discovery initiator packet to all of the plurality of processor units in the new node, the discovery initiator packet containing the connection information set up at the discovery initiator node for the new node; and

(d) receiving a discovery confirmation packet at the monitor process of the discovery initiator node, the discovery confirmation packet containing connection information enabling the discovery initiator node to directly communicate with the monitor process of the new node;

(e) sending, in response to the discovery confirmation packet, a discovery acknowledgement packet directly to the monitor process of the new node.

30. A method of automatically interconnecting a first new node and a second new node into a cluster of other nodes, the first new node having a lower node number than the second new node,

each node including a plurality of processor units and an internal switching fabric, one of the processor units in each node hosting a monitor process, comprising:

(a) broadcasting a discovery probe packet from the monitor process in the first new node to each of the plurality of processor units of the second new node and all the other nodes in the cluster, the discovery probe packet containing configuration information about the first new node;

(b) broadcasting, concurrently with step (a), a discovery probe packet from the monitor process in the second new node to each of the plurality of processor units of the first new node and all the other nodes in the cluster, the discovery probe packet containing configuration information about the second new node;

(c) receiving at least one discovery initiator packet at the monitor process of the first new node from each one of the other nodes acting as a first set of discovery initiator nodes, with each discovery initiator packet containing connection information enabling the first new node to directly communicate with the monitor processes of the first set of discovery initiator nodes;

(d) receiving at least one discovery initiator packet at the monitor process of the second new node from each one of the first new node and the other nodes acting as a second set of discovery initiator nodes, with each discovery initiator packet containing connection information enabling the second new node to directly communicate with the monitor processes of the second set of discovery initiator nodes;

(e) setting up, at each of the new nodes, in response to the discovery initiator packets, connection information enabling the discovery initiator nodes to directly communicate with the monitor process of each of the new nodes;

(f) sending a discovery confirmation packet by the first new node directly to the monitor processes of each one of the first set of discovery initiator nodes, with each discovery confirmation packet containing the connection information, set up at the first new node, and specific to each node of the first set of discovery initiator nodes to which the packet is sent;

(g) sending a discovery confirmation packet by the second new node directly to the monitor processes of each one of the second set of discovery initiator nodes, with each discovery confirmation packet containing the connection information, set up at the second new node, and specific to each node of the second set of discovery initiator nodes to which the packet is sent; and

(h) receiving, at each of the new nodes, a discovery acknowledgement packet in response to the discovery confirmation packet.

31. A method of automatically interconnecting a first new node and a second new node into a cluster of other nodes, each node having a node number and including a plurality of processor units and an internal switching fabric, one of the processor units in each node hosting a monitor process, comprising:

(a) receiving at least two discovery probe packets at the monitor processes of first new node, second new node and the other nodes in the cluster, one of the two discovery probe packets containing configuration information about the first new node, the other of the two discovery probe packets containing configuration information about the second new node, the first new node having a lower node number than the second new node;

(b) setting up, at each of the other nodes, in response to the discovery probe packets sent by the first new node, connection information enabling the first new node to directly communicate with the monitor processes of the other nodes;

(c) setting up, at the first new node and each of the other nodes, in response to the discovery probe packets sent by the second node, connection information enabling the second new node to directly communicate with the monitor processes of the first new node and the other nodes;

(d) sending discovery initiator packets from the other nodes to all of the plurality of processor units in the first new node, the discovery initiator packets containing the connection information set up at the other nodes for the first new node;

(e) sending discovery initiator packets from the first new node and the other nodes to all of the plurality of processor units in the second new node, the discovery initiator packets containing the connection information set up at first new node and the other nodes for the second new node; and

(f) receiving a discovery confirmation packet at the monitor process of each one of the other nodes, the discovery confirmation packet containing connection information enabling each one of the other nodes to directly communicate with the monitor process of the first new node;

(g) receiving a discovery confirmation packet at the monitor processes of each one of the first new node and the other nodes, the discovery confirmation packet containing connection

information enabling each one of the first new node and the other nodes to directly communicate with the monitor process of the second new node;

(h) sending, in response to each discovery confirmation packet received in step (f), a discovery acknowledgement packet from the other nodes directly to the monitor process of the first new node; and

(j) sending, in response to each discovery confirmation packet received in step (g), a discovery acknowledgement packet from the first new node and the other nodes directly to the monitor process of the second new node.